فصلنامه پژوهش‌های اقتصادی ایران/ سال بیست‌ویکم/ شماره ۶۶/ بهار ۱۳۹۵/ صفحات ۱۲۲-۱۰۳

مدل احتمال مشارکت زنان در ایران: الگوی لاجیت ناپارامتری

بهروز کشته‌گر[۱]
زینب سارانی[۲]

تاریخ ارسال: ۱۳۹۳/۹/۱۵ تاریخ پذیرش: ۱۳۹۵/۴/۶

چکیده

در این مقاله، یک روش مدل‌سازی لاجیت ناپارامتری برای تخمین احتمال مشارکت کار زنان ایران مطابق با داده‌های هزینه – درآمد خانوار سال ۱۳۸۸، معرفی شده است. تابع لجستیک برای احتمال مشارکت زنان به کمک روش حداکثر درست‌نمایی براساس متغیرهایی مانند موقعیت جغرافیایی (شهری/روستایی) درآمد شوهر، سواد زن، سن زن، درآمد غیرکاری، تعداد بچه بالای شش سال و زیر شش سال برازش شده است. دقت مدل‌سازی دو الگوی لاجیت پارامتری و ناپارامتری با استفاده از آماره‌هایی مانند وایت، ضریب اطمینان و میانگین مجذور خطا ارزیابی شده است. در انتها، اثر نهایی متغیرهای ورودی بر احتمال مشارکت زنان براساس ضرایب برازش شده مدل پارامتری و مدل ناپارامتری ارایه شده، برآورد شده است. نتایج حاکی از آن است که مدل‌سازی ناپارامتری دقت بیشتری نسبت به الگوی لاجیت پارامتری دارد. بیشترین تأثیر مثبت و منفی بر مشارکت زنان ایرانی به ترتیب براساس سواد زن و بچه زیر شش سال بوده است.

واژگان کلیدی: مدل لاجیت ناپارامتری، روش حداکثر درست‌نمایی، مدل‌سازی مشارکت زنان.
طبقه‌بندی JEL: *J21، J05، G10.*

۱- دکترای عمران، دانشکده مهندسی، دانشگاه زابل (نویسنده مسؤول)، پست الکترونیکی: Bkeshtegar@ uoz.ac.ir.
۲- دکترای اقتصاد دانشگاه تربیت مدرس، پست الکترونیکی: Elaha.shargh@yahoo.com.

۱- مقدمه

مشارکت زنان در فعالیت‌های اقتصادی نه‌تنها در توسعه ظرفیت‌ها و توانایی‌های زنان، بلکه در توسعه اقتصادی و اجتماعی جامعه نقشی کلیدی دارد. وقتی کشوری در مدار توسعه اقتصادی قرار می‌گیرد، این توسعه اقتصادی در کاهش شکاف مشارکت زنان و مردان در بسیاری از حوزه‌ها از جمله دسترسی به بازار کار و کاهش بیکاری مؤثر است. شاخص‌هایی مانند آموزش زنان، بارداری، امید به زندگی و درآمد سرانه نشان می‌دهند که ایران رشد قابل ملاحظه‌ای در مشارکت زنان نسبت به ده سال گذشته، داشته است (کولایی، ۱۳۸۳، ص ۹). مشارکت زنان در بازار کار موجب افزایش رقابت آزاد، بهره‌گیری کشور از بهترین استعدادها، افزایش بهره‌وری و رشد مستمر اقتصادی می‌شود.

طی سال‌های اخیر، غالباً در کنکور ورودی دانشگاه‌ها، دختران از پسران پیشی گرفته‌اند؛ برای مثال، در کنکور سال ۱۳۷۸، پذیرفته‌شدگان دختر ۵۷/۲ درصد کل پذیرفته‌شدگان را تشکیل می‌دادند[۱]. از سویی، با توجه به شرایط اقتصادی جامعه، نگرش عمومی نسبت به اشتغال و تحصیل دختران تغییر کرده است. در حال حاضر شاهد افزایش مشارکت زنان در بازار کار به علت کسب تحصیلات بالاتر هستیم (کولایی، ۱۳۸۳، ص ۲۸). ازاین‌رو، افزایش میزان تحصیلات زنان موجب کاهش نرخ بارداری شده، به‌طوری که این نرخ از ۷/۲ در سال ۱۳۶۵ به ۱/۹ در سال ۱۳۸۵ کاهش یافته است (موحدی، ۱۳۸۱). این، بدان معناست که بین میزان موالید و میزان اشتغال زنان در ایران رابطه‌ای منفی وجود داشته است. پایین‌تر بودن شاخص بارداری مبین آن است که زنان وقت کمتری را صرف نگهداری بچه می‌کنند. ازاین‌رو، زمان بیشتری را می‌توانند صرف فعالیت‌های درآمدزا کنند. از سویی، تربیت فرزندان زیر ۶ سال اثری منفی بر احتمال مشارکت زن در بازار کار می‌گذارد.

تخمین معادله مشارکت زنان این امکان را فراهم می‌آورد تا به بررسی پارامترهای مؤثر بر مشارکت زنان پرداخت. برای تخمین معادله مشارکت زنان که در آن متغیر وابسته

۱- مجله زنان، سال ۱۳۷۸، شماره ۵۶، ص ۶۸.

بیان‌کننده دو حالت کارکردن (یک) و کار نکردن (صفر) باشد، می‌توان از مدل احتمالی لاجیت استفاده کرد (بنکو و مارتینز[1] (۲۰۰۹) و معّلوف و ترافلیس[2] (۲۰۱۱)). مدل لاجیت پارامتری برای زنان با مشخصات فردی و اجتماعی یکسان، اما عکس‌العمل‌های متفاوت، یک ضریب برای تابع لجستیک برآورد می‌کند. ازاین‌رو، این روش قادر به مدل‌سازی صحیح پدیده‌های دودویی مقطعی نیست. از سوی دیگر، مشکل رایج داده‌های دودویی در مدل‌سازی لاجیت پارامتری وجود ناهمسانی واریانس است (فرولیک[3] (۲۰۰۶) و معّلوف و ترافلیس (۲۰۰۸)). برای رفع ناهمسانی واریانس به تکنیک‌های مدل‌سازی ناپارامتری با یک تابع کرنل (فرولیک (۲۰۰۶) و معّلوف و ترافلیس (۲۰۰۸)) یا الگوی مدل‌سازی غیرخطی (سارانی و همکاران، ۱۳۹۳) روی آورده‌اند. ازاین‌رو، رفع ناهمسانی واریانس و دستیابی به یک مدل با خطای کمتر می‌تواند در مدل‌سازی لاجیت مهم باشد.

در این مقاله، یک مدل ناپارامتری لاجیت بیان شده است که می‌تواند خطای مدل پیش‌بینی مشارکت کار زنان ایرانی را نسبت به مدل پارامتری کاهش دهد. این مدل با استفاده از یک بخش جریمه در برآوردگر حداکثر درست‌نمایی دودویی برآورد می‌شود. براساس ضرایب حاصل الگوی پارامتری و ناپارامتری مدل‌سازی لاجیت، تأثیر متغیرهای ورودی تخمین زده شده است. برای دستیابی به رفع ناهمسانی واریانس و تخمین اثر نهایی متغیرهای ورودی بر احتمال مشارکت زنان، مقاله حاضر در هفت بخش تدوین شده است؛ در بخش دوم، به بیان ادبیات موضوع می‌پردازیم. مبانی نظری مدل‌سازی دودویی به کمک روش حداکثر درست‌نمایی در بخش سوم، ارایه می‌شود. در بخش چهارم، داده‌های هزینه- درآمد خانوار بیان می‌شود. بخش پنجم به نتایج مدل‌سازی پارامتری و ناپارامتری براساس معیارهای اقتصادسنجی و چندین آماره مانند ضریب اطمینان و میانگین مجذور خطا اختصاص دارد. بخش ششم، تأثیر متغیرهای توصیفی براساس ضرایب حاصل از مدل‌سازی پارامتری و ناپارامتری برای میزان مشارکت زنان ارایه می‌شود و اثرات متغیر

1- Bianco and Martínez
2- Maalouf and Trafalis
3- Frolich

تحصیلات و تعداد بچه زیر شش سال برمبنای مدل‌سازی پارامتری و ناپارامتری مورد بحث قرار می‌گیرد. در انتها، خلاصه‌ای از نتایج ارایه می‌شود.

۲- ادبیات موضوع

میزان مشارکت اعضای خانوار در بازار کار یکی از موضوع‌های مهم در اقتصاد است. کلینگورس[۱] (۱۹۸۶)، برای ارزیابی میزان مشارکت اعضای خانوار، یک مدل ریاضی را براساس ارتباط اعضای خانواده برای تصمیم‌گیری در بازار کار توسعه داد. همچنین چیاپوری[۲] (۱۹۸۸ و ۱۹۹۲)، مدل عرضه کار جمعی را با استفاده از یک تابع مطلوبیت منحصربه‌فرد پیشنهاد کرد. دونی[۳] (۲۰۰۷)، مدل عرضه کار ارایه شده توسط چیاپوری (۱۹۸۸ و ۱۹۹۲) را با تابع نیمه‌لگاریتمی بیان کرد. برولاوا و چیکاوا[۴] (۲۰۱۲)، مدل‌های عرضه کار خانوار را برای کشورهای رومانی، گرجستان و فرانسه با استفاده الگوی چیاپوری و همکاران (۱۹۹۲)، با یکدیگر مقایسه کردند.

لکنر[۵] (۱۹۹۱)، با استفاده از مدل لاجیت، به بررسی مشارکت در بازار کار زنان فرانسه پرداخت. وی نشان داد که سن و تحصیلات تأثیر مثبت و تعداد بچه زیر شش سال و درآمد همسر اثر منفی بر مشارکت زن دارد. بعدها، گرفین[۶] (۱۹۹۶)، به کمک مدل پروبیت، مشارکت زنان متأهل آلمانی و سوئدی را برآورد و با مدل پارامتری لکنر (۱۹۹۱) مقایسه کرد. سپس، براساس یک مدل شبه‌پارامتری ارایه شده توسط کلین و اسپادی[۷] (۱۹۹۳)، به رفع ناهمسانی واریانس اقدام کرد و نتیجه گرفت که تحصیلات زن و تعداد فرزندان به ترتیب اثر مثبت و منفی بر مشارکت وی دارد. عرضه کار زنان متأهل در مکزیک توسط گونگ و سوئست[۸] (۲۰۰۰) بررسی شد. همچنین گارسیا[۹] و همکاران (۲۰۰۲)، به بررسی

1- Killingworth
2- Chiappori
3- Donni
4- Berulava and Chikava
5- Lechner
6- Gerfin
7- Klein and Spady
8- Gong and Soest
9- García

رفتار نیروی کار زنان اسپانیا پرداختند و نتیجه گرفتند که درآمد خانوار و تعداد فرزندان اثر منفی بر احتمال حضور زنان اسپانیا دارد. سارانی و همکاران (۱۳۹۳)، براساس یک مدل لاجیت غیرخطی و کشاورز حداد و باقری قنبرآبادی (۱۳۹۰)، با استفاده از مدل‌سازی لاجیت ناپارامتریک با تابع کرنل هیبریدی نشان دادند که بچه زیر شش سال و تحصیلات به ترتیب اثرات منفی و مثبت بر مشارکت زنان ایرانی دارند. همچنین کشاورز حداد و علویان قوانینی (۱۳۹۱)، نتیجه گرفتند که با افزایش میزان تحصیلات، شکاف جنسیتی بین زنان و مردان ناچیز می‌شود.

فرولیک (۲۰۰۶)، رگرسیون ناپارامتری با تابع کرنل هیبریدی (کارول[1] و همکاران، ۱۹۹۸) را راهکاری مناسب برای رفع ناهمسانی واریانس در مدل‌سازی لاجیت معرفی کرد. کشاورز حداد و باقری قنبرآبادی (۱۳۹۰)، براساس الگوی لاجیت ناپارامتری (فرولیک، ۲۰۰۶)، ناهمسانی واریانس مدل پیش‌بینی مشارکت زنان ایران را رفع کردند. معّلوف و ترافلیس (۲۰۱۱)، برای برآورد داده‌های دودویی، روش مدل‌سازی لاجیت وزنی را با اضافه کردن یک بخش جریمه تنظیم کردند. مدل‌سازی لاجیت ناپارامتری تعمیم‌یافته در تخمین داده‌های دودویی ساده‌تر از مدل‌سازی لاجیت ناپارامتری با تابع کرنل هیبریدی و مدل لاجیت غیرخطی است. ازاین‌رو، مدل‌سازی تنظیم شده لاجیت ناپارامتری را می‌توان به‌سادگی برای برآورد احتمال مشارکت زنان به کار گرفت و براساس آن، تأثیر متغیرهای توضیحی را بر میزان مشارکت زنان تخمین زد. فرضیه‌های اثر مثبت تحصیلات زن و اثر منفی بچه زیر شش سال بر میزان مشارکت زنان را می‌توان با استفاده از یک الگوی تحلیل حساسیت ارایه شده در این مقاله، برمبنای مدل لاجیت پارامتری و ناپارامتری ارزیابی کرد.

۳- مبانی نظری مدل‌سازی مشارکت زنان

نکته مهم در بحث مشارکت نیروی کار، توجه به حداقل دستمزد[2] (w^*) است. یک فرد ساعتی را که صرف استراحت می‌کند با محدودیت زمانی به صورت $T = L + H$ مواجه

1- Carroll
2- Reservation Wage

بوده که در آن، T کل زمان موجود، H ساعات کاری که فرد به‌اندازه w دستمزد گرفته و L زمان فراغت است. ازاین‌رو، محدودیت بودجه فرد به صورت زیر قابل بیان است:

$$wT + A = wL + G \qquad (۱)$$

رابطه یادشده نشان می‌دهد که محدودیت بودجه شامل کل زمان در دسترس است که با نرخ دستمزد فرد برآورد شده (wT) و ثروت (A) به آن اضافه می‌شود. کل زمان صرف استراحت یا هزینه فراغت (wL) وکالای مصرفی فرد (G) می‌شود. برای فردی که کار کردن را بر فراغت ترجیح می‌دهد، حداقل دستمزد مطلوب (w^*) بزرگ‌تر از دستمزد حداقل است ($w^* > w_0$). اگر معادله دریافتی دستمزد فرد به صورت زیر باشد:

$$w_i = \exp(X_{1i}\beta_1 + u_{1i}) \qquad (۲)$$

که در آن، X_{1i} بردار برخی از ویژگی‌های فردی و اجتماعی فرد iام است، آنگاه می‌توان گفت، افرادی که تمایل به کار کردن ندارند، دستمزدی برابر با صفر ($w=0$) دارند، بنابراین، حداقل دستمزد مطلوب افرادی که تمایل به کار کردن دارند، تابعی از ترجیحات فردی آنهاست:

$$w_i^* = \exp(X_{2i}\beta_2 + A_i\gamma_2 + u_{2i}) \qquad (۳)$$

که در آن، A بردار ثروت است. متغیر X_2 بردار ویژگی فردی و اجتماعی بوده که ممکن است شامل مؤلفه‌های مشترک با متغیر X_1 نیز باشد و u_1 و u_2 جزء خطاست. با لگاریتم‌گیری معادله دستمزد فرد (رابطه ۲) و معادله حداقل دستمزد (رابطه ۳) داریم:

$$Ln(w_i) = X_{1i}\beta_1 + u_{1i} \qquad (۴)$$

$$Ln(w_i^*) = X_{2i}\beta_2 + A_i\gamma_2 + u_{2i} \qquad (۵)$$

برای افرادی که کار دارند، دستمزد تحقق‌یافته آنها غیرصفر است ($Ln(w_i) > 0$). ازاین‌رو، افرادی که تمایل به کار کردن در بازار کار دارند $Ln(w_i) \geq Ln(w_i^*)$ که با تفاضل دو رابطه ۴ و ۵ داریم:

$$Ln(w_i) - Ln(w_i^*) = X_{1i}\beta_1 - X_{2i}\beta_2 - A_i\gamma_2 + \varepsilon_i > 0 \qquad (۶)$$

با توجه به معادله یادشده مشاهده می‌شود که تمایل به مشارکت در بازار کار تابعی از مشخصات فردی و اجتماعی زنان است، ازاین‌رو، می‌توان گفت:

$$y_i = \begin{cases} 1 & Ln(w_i) > Ln(w_i^*) \\ 0 & Ln(w_i) \le Ln(w_i^*) \end{cases} \qquad (۷)$$

بنابراین، احتمال حضور زنان در بازار کار را می‌توان به صورت زیر تعریف کرد:

$$\Pr(y_i = 1 | Xi, \beta) = 1 - F(-\beta' X) = F(\beta' X) \qquad (۸)$$

که در آن، F تابع توزیع تجمعی لاجستیک است (معلوف و ترافلیس، ۲۰۱۱).

$$F(-\beta' X) = pi = \frac{\exp[f(\beta xi)]}{1+\exp[f(\beta xi)]} = \frac{1}{1+\exp[f(\beta xi)]} \qquad (۹)$$

که در آن، $f(\beta x)$ مقدار تابع پیشگویی است که به صورت زیر بیان می‌شود:

$$f(\beta x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + ... + \beta_n x_n \qquad (۱۰)$$

در تابع یادشده هر مشاهده با توزیع برنولی ($F(-\beta' X)$) براساس ضرایب (β) و متغیرهای مستقل فردی و اجتماعی زنان (X) شناخته می‌شود. ازاین‌رو، الگوی پارامتری حداکثر درست‌نمایی برای تخمین احتمال مشارکت زنان به صورت زیر بیان می‌شود:

$$ML(\beta) = \sum_{i=1}^{n} \ln[pi] \;\; yi + \ln[1 - pi](1 - yi) \qquad (۱۱)$$

مدل لاجیت پارامتری برای داده‌های با اثر مقطعی ایجاد ناهمسانی واریانس می‌کند (کشاورز حداد و باقری قنبرآبادی، ۱۳۹۰). مدل لاجیت پارامتری برای تخمین پارامترهای تابع مشارکت زنان برپایه فروض استاندارد کلاسیک استوار است. در صورت برقرار نبودن این فروض، تخمین حاصل از اعتبار کمتری برخوردار است. برای پرهیز از نقص‌های مدل لاجیت از تکنیک‌های اقتصادسنجی ناپارامتری استفاده شده است (فرولیک (۲۰۰۶) و کشاورز حداد و قنبرآبادی (۱۳۹۰)) که بتوان به کمک آن فرض تصادفی بودن تخمین را برای هر نفر به صورت مجزا لحاظ کرد. تابع تخمین‌زن حداکثر درست‌نمایی برای روش‌های لاجیت ناپارامتری به صورت زیر است:

$$ML(\beta) = \sum_{i=1}^{n} \ln[pi] \;\; yi + \ln[1 - pi](1 - yi) - \lambda \|\beta\|^2 \qquad (۱۲)$$

که در آن، $f(\beta xi)$ یک تابع از متغیرهای مستقل مسأله است. همچنین در معادله (۱۲) جزء جریمه $\lambda\|\beta\|^2$ برای به‌دست آمدن تنظیم مناسب ضرایب اضافه شده است. براساس الگوی ناپارامتری در معادله (۱۲)، جزء جریمه ($\lambda\|\beta\|^2$) می‌تواند انعطاف‌پذیری لازم را در برآورد مدل‌های دودویی فراهم کند.

٤- داده‌های هزینه- درآمد خانوار

برای تعیین مدل مشارکت زنان ایران از داده‌های هزینه- درآمد خانوار مرکز آمار ایران، گزارش سال ۱۳۸۸، استفاده شده است. داده‌های آماری مربوط به اطلاعات ۱۰۲۷۳ زن متأهل با سنین ۱۴ تا ۶۰ سال در نظر گرفته شده است. در این مجموعه، ۵۲۲۱ خانوار در شهر سکونت داشته‌اند و ۱۶۳۲ زن متأهل مشغول به کار بوده‌اند. خلاصه‌ای از خصوصیات آماری متغیرهای توضیحی در جدول شماره ۱، ارایه شده است.

تعداد فرزندان زیر ۶ سال ($UK6$) و بالای ۶ سال ($AK6$) هستند. متغیر شهری و روستایی (RU) مشخص‌کننده محل سکونت زن (شهر(عدد یک)، روستا (عدد صفر)) است، متغیر میزان تحصیلات زن (edu) است که به صورت ۶ رتبه، دسته‌بندی شده و عبارت است از: صفر بی‌سواد، عدد یک تحصیلات تا سوم راهنمایی، عدد ۲ تحصیلات بالای سوم راهنمایی و زیر دیپلم، عدد ۳ تحصیلات بالای دیپلم و زیر فوق دیپلم، عدد ۴ تحصیلات بالای فوق دیپلم و زیر لیسانس (دانشجویان دوره کارشناسی نیز در این رتبه قرار می‌گیرند)، عدد ۵ فوق لیسانس به بالا و پزشکی. متغیرهای ثروت (INC) است و دستمزد ماهیانه مرد (WM) برحسب یک میلیون ریال با تابع لگاریتمی در مدل‌سازی لحاظ شده است.

جدول ۱- خصوصیات آماری متغیرهای ورودی مدل

متغیرها	مشاهدات	میانگین	انحراف معیار	حداقل	حداکثر
تحصیلات	۱۰۲۷۳	۰/۹۴۴۶۱	۰/۹۳۶۷۴	۰	۵
شهری و روستایی	۱۰۲۷۳	۰/۵۰۸۲۳	۰/۴۹۹۹۶	۰	۱
دستمزد مرد	۱۰۲۷۳	۲/۹۰۹۱۰	۴/۲۳۴۳۵	۰	۱۶۶/۷
ثروت	۱۰۲۷۳	۵/۱۴۲۰۷	۱۹/۱۷۸۶۵	۰	۱۰۰۰
سن زن	۱۰۲۷۳	۳۵/۴۵۱۹۶	۱۰/۳۵۱۲۰	۱۴	۶۰
بچه زیر ۶ سال	۱۰۲۷۳	۰/۴۵۹۳۶	۰/۶۷۲۱۹	۰	۴
بچه بالای ۶ سال	۱۰۲۷۳	۱/۴۹۳۲۴	۰/۳۵۹۵۳	۰	۱۰

مأخذ: مرکز آمار ایران، اطلاعات هزینه - درآمد خانوار، سال ۱۳۸۸.

مطابق با روابط ۶ و ۷، کار کردن/ نکردن زن علاوه بر دستمزد مطلوب، تابعی از ترجیحات فردی نیز هست. وجود بچه زیر ۶ سال موجب افزایش حضور زن در خانه می‌شود، اما ترجیحات فردی مانند موقعیت جغرافیایی و سطح تحصیلات ممکن است موجب افزایش حضور زن در بازار کار شود. همچنین ثروت و درآمد همسر موجب می‌شود فرد زمان بیشتری را به استراحت اختصاص دهد، اما ترجیحات فردی مانند تحصیلات و سن تمایل فرد را برای حضور در بازار کار افزایش می‌دهد. بنابراین، بچه زیر شش سال، ثروت و درآمد همسر جزء متغیرهای درون‌زا به شمار می‌آیند.

۵- مقایسه نتایج مدل‌سازی لاجیت پارامتری و ناپارامتری

نتایج مدل‌سازی پارامتری و ناپارامتری براساس چندین ضریب تعمیم (λ) به کمک آماره‌های زیر مقایسه شده است.

۵ -۱- آماره‌های مقایسه مدل‌های پارامتری و ناپارامتری

۵ -۱-۱- جذر میانگین مربع خطا

جذر میانگین مربع خطا[1] ($RMSE$) به صورت زیر قابل محاسبه است:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y} - y_i)^2}{n}} \qquad (۱۳)$$

1- Root Mean Square Error (RMSE)

مقدار $RMSE$ کمتر بیان‌کننده برآورد صحیح و انطباق بیشتر داده‌های پیش‌بینی شده بر مشاهده شده است.

۵-۱-۲- ضریب اطمینان[1]

ضریب اطمینان براساس حاصل‌ضرب دو معیار ضریب تطبیق[2] و ضریب کارآیی[3] به صورت زیر قابل محاسبه است (ویلموت و ماتسورا[4]، ۱۹۹۸):

$$CI = D \times EF \qquad (۱۴)$$

که در آن، D ضریب تطبیق و EF ضریب کارآیی است، ازاین‌رو، داریم:

$$D = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(|\hat{y}_i - \bar{\hat{y}}| + |y_i - \bar{y}|)^2} \qquad (۱۵)$$

$$EF = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\bar{y} - y_i)^2} \qquad 0 \le EF \le 1 \qquad (۱۶)$$

که در آن، $\bar{y}$ و $\bar{\hat{y}}$ به ترتیب متوسط مشارکت مشاهده شده و مشارکت پیش‌بینی شده است. y_i و $\hat{y}_i$ به ترتیب مقدار پیش‌بینی احتمال مشارکت و مشاهدات مشارکت خانوار iام است. CI=۰ و CI=۱ *به ترتیب* دلالت بر عدم وابستگی و تطبیق کامل داده‌های پیش‌بینی شده و مشاهدات دارد (کشته‌گر[5] و همکاران، ۲۰۱۶).

۵ -۱-۳- تعداد یک‌های تخمین زده شده

میزان برآورد تعداد یک‌ها (زنان شاغل) برای هر مدل براساس احتمال بیشتر از ۰/۵ به صورت زیر محاسبه شده است:

1- Confidence Index (CI)
2- Agreement Index
3- Efficient Factor (EF)
4- Willmott and Matsuura
5- Keshtegar

$$P_i = \begin{cases} 1 & if\ \hat{y}_i \geq 0.5 \\ 0 & if\ \hat{y}_i < 0.5 \end{cases} \qquad (۱۷)$$

که در آن، P_i مقدار احتمال مشارکت زن به صورت صفر (عدم مشارکت در بازار کار) و یک (مشارکت در بازار کار) است.

۵-۲- مقایسه مدل ناپارامتری و پارامتری

در جدول شماره ۲، مقادیر ضرایب مدل پارامتری ($\lambda = 0$) و مدل ناپارامتری به‌ازای ضرایب تنظیم مختلف ارایه شده که این ضرایب برای بخش‌های مختلف متغیرهای توضیحی معادله مشارکت به صورت زیر است:

$$f(X\beta) = \beta_0 + \beta_1 UR + \beta_2 age + \beta_3 \log(WM) + \beta_4 \log(INC) + \\ + \beta_5 UK6 + \beta_6 AK6 + \beta_7 edu \qquad (۱۸)$$

همچنین در جدول شماره ۲، آماره‌های مقایسه که در بخش ۵-۱ ارایه شد، آمده است. مشخص است که با تغییر ضریب تنظیم مدل ناپارامتری (λ) نه‌تنها ضرایب برآورد شده تغییر ناچیزی داشته، بلکه آمارهای مقایسه نیز تغییرات ناچیزی را نشان داده است. مقدار $RMSE$ در مدل‌های ارایه شده با ضرایب تنظیم مختلف، نسبت به مدل پارامتری مقدار کمتری را نشان داده است. همچنین مشخص است که ضریب برآورد شده بچه زیر شش سال در مدل‌های ناپارامتری تغییر علامت (از علامت مثبت به منفی) داده است. افزایش زمان حضور زن در خانه به علت نقش تربیتی مادر و نیاز به نگهداری کودک زیر شش سال، به کاهش احتمال حضور زن در بازار کار منجر می‌شود. این، بدان معناست که وجود بچه زیر شش سال در یک خانوار بر مشارکت زن اثر منفی دارد. ازاین‌رو، می‌توان استنباط کرد که روش مدل‌سازی ناپارامتری ارایه شده، توانمندی تطبیق و ارایه مناسب نتایج را براساس داده‌های ارایه شده، نسبت به مدل پارامتری دارد و همچنین موجب کاهش خطای حاصل نسبت به مدل پارامتری شده است.

مدل ناپارامتری اثر مثبت بیشتری از متغیر تحصیلات زن (*edu*) را نسبت به مدل پارامتری نشان می‌دهد و این موضوع سبب تأثیر مثبت معنادار تحصیلات زنان بر میزان احتمال مشارکت آنان در بازار کار می‌شود.

در مدل‌های ناپارامتری، ضریب اطمینان نسبت به مدل پارامتری افزایش یافته است. مشخص است که در تمام مدل‌های ناپارامتری، تعداد یک‌های برآورد شده بیشتر از مدل پارامتری است. با توجه به در نظر گرفتن تعداد یک‌ها، اثر تطبیق داده‌ها و خطای حاصل از مدل‌سازی مشخص است که به‌ازای ضریب تنظیم مدل ناپارامتری برابر با ۱۵۰، نتایج مناسبی از مدل‌سازی ارایه شده است.

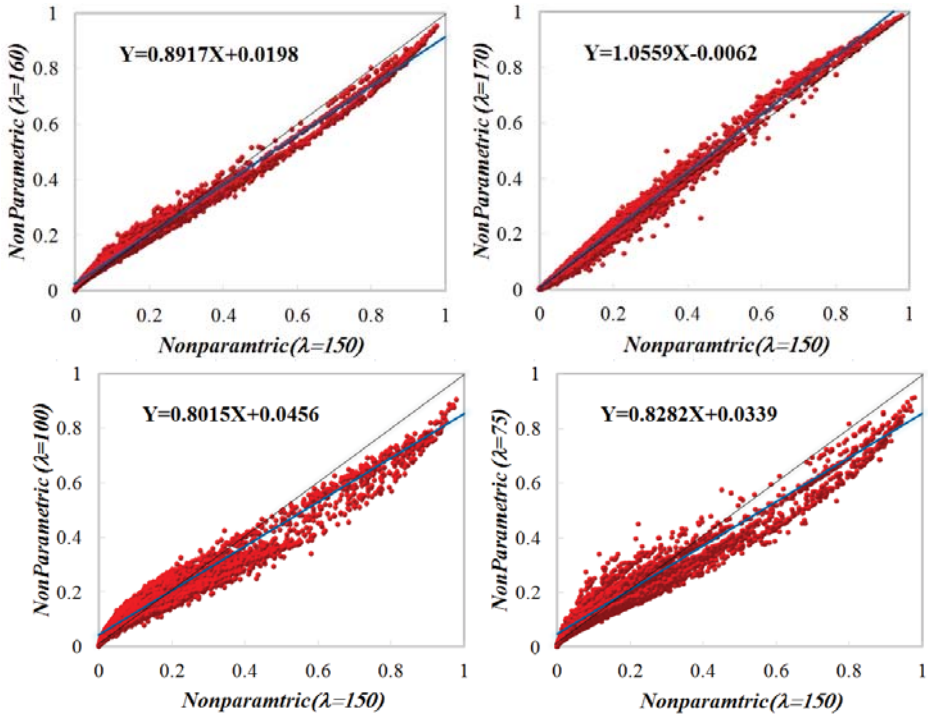
جدول ۲- آماره‌های مقایسه مدل لاجیت پارامتری و ناپارامتری

ضریب تنظیم	۰ (پارامتری)	۱۲۵	۱۵۰	۱۷۰
β_0	-۲/۶۸۳۸	-۵/۵۷۸۶	-۵/۹۸۴۵	-۶/۰۳۰۸
UR	-۰/۰۲۱۹	۰/۰۹۹۷	۰/۰۹۶۴	۰/۱۰۱۶
age	۰/۰۱۱۱	۰/۰۶۳۳	۰/۰۶۸۰	۰/۰۶۶۸
log(WM)	-۰/۰۶۰۸	-۰/۰۸۳۸	-۰/۰۸۷۵	-۰/۹۲۲
log(INC)	-۰/۰۰۳۴	-۰/۰۰۶۱	-۰/۰۰۴۲	-۰/۰۰۵۰
UK6	۰/۰۱۶۶	-۰/۳۶۶۷	-۰/۳۹۱۰	-۰/۳۱۵۰
AK6	۰/۰۶۷۷	۰/۲۰۲۰	۰/۲۱۲۹	۰/۲۸۱۰
edu	۰/۶۴۸۶	۱/۱۰۳۵	۱/۱۵۹۵	۱/۲۸۶۹
آماره‌های مقایسه مدل‌سازی				
RMSE	۰/۳۵۱۲۱	۰/۳۵۰۹۸	۰/۳۵۰۹۵	۰/۳۵۱۰۳
CI	۰/۰۲۴۲	۰/۰۲۹۹	۰/۰۳۳۱	۰/۰۲۹۵
تعداد یک‌ها	۲۳۳	۴۱۲	۴۴۷	۵۰۸

مأخذ: محاسبات پژوهش.

در شکل شماره ۱، نمودار پراکندگی چندین مدل ناپارامتری با ضرایب تنظیم مختلف ($\lambda = 75,100,160,170$) نسبت به مدل ناپارامتری با ضریب تنظیم ۱۵۰ ($\lambda = 150$) نشان داده شده است. همچنین در این شکل، معادله خط پراکندگی داده‌های مدل‌های ناپارامتری

نسبت به مدل ناپارامتری با λ=150 رسم شده است. مشخص است که پراکندگی داده‌های پیش‌بینی مدل ناپارامتری با ضریب λ=150 نسبت به مدل‌های ناپارامتری با ضرایب λ=160,170 چندان نیست و نمی‌توان تمایزی بین آنها قایل شد. می‌توان اظهار داشت که حساسیت مدل‌سازی به تغییرات متغیرها به‌ازای $150 \leq \lambda \leq 170$ کاهش یافته است.



شکل ۱- مقایسه نمودار پراکندگی مدل‌های ناپارامتری

٦- تأثیر متغیرهای توضیحی بر مشارکت زنان ایرانی

براساس متغیرهای محاسبه شده مطابق با مدل‌سازی لاجیت پارامتری و ناپارامتری (نتایج ارایه شده در جدول شماره ۲)، تأثیر هر متغیر توضیحی براساس رابطه پیشنهادی (۱۹) بر میزان احتمال مشارکت زنان سنجیده شده است.

$$\alpha_{Xi} = \frac{\beta_{Xi}^{NP} - \beta_{Xi}^{P}}{\sqrt{\sum_{i=1}^{N}(\beta_{Xi}^{NP} - \beta_{Xi}^{P})^2}} \qquad (۱۹)$$

که در آن، α_{Xi} مقدار حساسیت یا تأثیر متغیر مستقل ورودی Xi است که براساس اختلاف بین ضرایب حاصل برای متغیر مورد نظر در مدل‌سازی لاجیت ناپارامتری (β_{Xi}^{NP}) و پارامتری (β_{Xi}^{P}) حاصل شده است. با توجه به رابطه یادشده و نتایج جدول شماره ۲، میزان تأثیر هر متغیر توضیحی بر مشارکت زنان ایرانی براساس برآوردگر لاجیت پارامتری و ناپارامتری ارایه شده در این مقاله، در جدول شماره ۳، نشان داده شده است. در جدول شماره ۳، به‌ازای سه ضریب تنظیم مختلف در مدل‌سازی ناپارامتری، مقدار حساسیت هر متغیر مطابق با رابطه (۱۹) ارایه شده است.

جدول ۳- مقایسه مقدار اثر (حساسیت) متغیرهای توضیحی بر میزان مشارکت زنان متأهل ایرانی

ضریب تنظیم	۱۲۵	۱۵۰	۱۷۰
UR	۰/۱۹۷۴	۰/۱۷۳۲	۰/۱۷۵۱
age	۰/۰۸۳۷	۰/۰۸۳۴	۰/۰۷۹۰
Log(WM)	-۰/۰۳۶۹	-۰/۰۳۹۲	-۰/۰۴۴۶
Log(INC)	-۰/۰۰۴۴	-۰/۰۰۱۱	-۰/۰۰۲۳
UK6	-۰/۶۱۳۸	-۰/۵۹۷۰	-۰/۴۹۷۱
AK6	۰/۲۱۵۱	۰/۲۱۲۸	۰/۳۰۲۳
edu	۰/۷۲۸۵	۰/۷۴۸۲	۰/۷۸۳۳

مأخذ: داده‌های پژوهش.

همان‌طور که مشخص است، بیشترین تأثیر مثبت مربوط به متغیر سواد و بیشترین تأثیر منفی به‌ازای بچه زیر ۶ سال حاصل شده است. نتایج اثرات مثبت و منفی به‌ازای تمام ضرایب تنظیم شده است، به‌طوری که متغیرهایی مانند شهری و روستایی، سن، بچه بالای ۶ سال و تحصیلات اثر مثبت را بر مشارکت و متغیرهایی از جمله: درآمد همسر، ثروت و بچه زیر ۶ سال اثر منفی را بر احتمال مشارکت زنان ایرانی نشان داده است. درآمد همسر و

ثروت اثر کمتری را نسبت به سایر متغیرها بر کار کردن زنان در بازار کار نشان داده است. همچنین مشخص بوده که کمترین اثر مثبت به‌ازای سن زن حاصل شده است.

تبدیل موقعیت زندگی یک زن از روستا به شهر، موجب افزایش مشارکت وی در بازار کار شده است (به صورت اثر مثبت شهری و روستایی بر مشارکت). ممکن است وجود مراکز نگهداری کودک و همچنین دسترسی به امکانات شغلی بیشتر در مراکز شهری برای زنان شهری فرصت اشتغال بیشتری را نسبت به زنان روستایی فراهم کند.

افزایش ثروت موجب افزایش بودجه خانوار می‌شود، ازاین‌رو، توانایی فرد برای انجام فعالیت‌های رفاهی افزایش می‌یابد.این موضوع، ممکن است موجب کاهش احتمال مشارکت زنان در بازار کار شود. همچنین ثروت به‌عنوان درآمد موقتی برای خانوار شناخته می‌شود و درآمد موقتی نقشی در تصمیم‌گیری مشارکت زنان ایفا نمی‌کند. ازاین‌رو، این روش تحلیل حساسیت کمترین تأثیر منفی را به‌ازای ثروت نشان داده است.

لحاظ کردن بخش تنظیم ($\lambda\|\beta\|^2$) در مدل‌سازی ناپارامتری، موجب کاهش خطا و انطباق بیشتر داده‌های پیش‌بینی بر مشاهدات شده است. از سویی، تنظیم ضرایب برمبنای میزان حساسیت مناسب در مدل در نظر گرفته شده است، به‌طوری که روش محاسبه حساسیت متغیرهای ورودی بر کار کردن زن، فرض اثر مثبت تحصیلات و اثر منفی بچه زیر ۶ سال را تأیید می‌کند. همچنین این روش مدل‌سازی و تحلیل حساسیت حاصل از آن نشان داده است که سن تأثیر مثبت کمی بر مشارکت زن دارد که این نتایج مطابق با نتایج ارایه شده توسط کشاورز حداد و میر باقری (۱۳۹۰)، است.

یکی از آزمون‌های مناسب برای آزمایش کردن ناهمسانی واریانس آماره وایت است. مطابق با مدل ناپارامتری (ضریب تعمیم برابر با ۱۵۰) و مدل پارامتری مقدار آماره وایت به ترتیب برابر با ۴/۲۴ و ۳۴/ ۷۸ حاصل شده است. این موضوع نشان‌دهنده آن است که مدل ناپارامتری ناهمسانی واریانس را مرتفع کرده است. از سوی دیگر، ضریب کارآیی مطابق با رابطه (۱۶) برای دو مدل پارامتری و ناپارامتری به ترتیب برابر با ۰/۰۷۷ و ۰/۱۲۵ حاصل شده که انطباق مناسب داده‌های مدل ناپارامتری بر مشاهدات را نشان می‌دهد. مطابق با

آماره ضریب لاگرانژ مورد استفاده توسط براء و بلیس[1] (۲۰۰۱)، خروج دو متغیر ثروت و درآمد همسر بر همسانی واریانس در دو مدل پارامتری و ناپارامتری آزمون شده است. مقدار آماره ضریب لاگرانژ برای مدل پارامتری حدود ۴/۱۸ و برای مدل ناپارامتری ۰/۸۳ حاصل شده که با سطح معناداری ۰/۰۵ (آماره کای دو ۳/۸۴) مقدار آماره مدل ناپارامتریک در محدوده مجاز قرار گرفته است. ازاین‌رو، مدل پارامتری نسبت به متغیرهای ثروت و درآمد همسر ناهمسانی واریانس دارد. همچنین متغیرهای ثروت و بچه بالای ۶ سال نیز در مدل پارامتری موجب ناهمسانی واریانس خطا شده (ضریب لاگرانژ حاصل برابر۱۳/۶۴)، اما مدل ناپارامتری ناهمسانی واریانس خطا را نشان نداده است (ضریب لاگرانژ برابر ۱/۱۸).

۷- خلاصه و نتیجه‌گیری

در این مقاله، احتمال مشارکت زنان ایرانی (شهری و روستایی) براساس متغیرهای مستقل پیوسته (دستمزد همسر و ثروت)، متغیرهای گسسته (سن، موقعیت جغرافیایی (شهری وروستایی) و تعداد فرزندان بالای ۶ سال و زیر ۶ سال) و متغیر رتبه‌بندی شده (تحصیلات) ارزیابی شده است. ضرایب مدل پیش‌بینی مشارکت زنان ایرانی براساس تخمین‌زن حداکثر درست‌نمایی لاجیت پارامتری و ناپارامتری برآورد شده است. می‌توان نتایج زیر را برای مدل‌سازی پارامتری و ناپارامتری بیان کرد:

- ✓ نتایج متفاوتی از ضرایب مدل مشارکت برمبنای برآوردگر لاجیت پارامتری و ناپارامتری حاصل شده است. ازاین‌رو، برآورد متفاوتی از احتمال مشارکت زنان در مدل‌سازی پارامتری و ناپارامتری به‌دست می‌آید.
- ✓ تأثیر بچه زیر ۶ سال در مدل پارامتری مثبت، اما در مدل ناپارامتری منفی بوده است. ازاین‌رو، مدل ناپارامتری ارایه شده، پیش‌بینی صحیح‌تری نسبت به مدل پارامتری برای احتمال مشارکت زنان ایرانی نسبت به حضور بچه زیر ۶ سال نشان داده است.

1- Bera and Bilias

- ✓ ضریب مثبت تحصیلات در مدل‌سازی ناپارامتری بیشتر از مدل‌سازی پارامتریک محاسبه شده است.
- ✓ روش مدل‌سازی ناپارامتری به‌سادگی روش پارامتری بوده، اما نسبت به مدل‌سازی پارامتری، انطباق بیشتری را نشان داده و علاوه بر کاهش خطا، ناهمسانی واریانس را نیز مرتفع کرده است.

در این مقاله، یک فرآیند محاسبه تأثیر متغیرهای توضیحی بر مدل احتمال مشارکت زنان ارایه شده است که برمبنای آن، اثر نهایی هر متغیر با استفاده از ضرایب برآورد شده مدل ناپارامتری و پارامتری قابل محاسبه است. برمبنای تحلیل حساسیت ارایه شده، می‌توان نتایج زیر را به‌طور خلاصه بیان کرد:

- ✓ بیشترین اثر مثبت بر احتمال مشارکت زنان برمبنای متغیر تحصیلات حاصل شده است.
- ✓ متغیر بچه زیر ۶ سال اثر منفی چشمگیری بر احتمال حضور زن در بازار کار داشته است.
- ✓ با توجه به افزایش ثروت خانوار و میزان استفاده از ساعات فراغت بیشتر، ممکن است احتمال حضور زن در بازار کار کاهش یابد. مطابق با روش مدل‌سازی ارایه شده، متغیرهای ثروت و درآمد همسر تأثیر منفی بر کار کردن زن داشته است.
- ✓ کمترین تأثیر مثبت و منفی بر احتمال مشارکت زنان ایرانی به ترتیب با توجه به متغیر سن و ثروت حاصل شده است.

منابع

مرکز آمار ایران، *سالنامه آماری کشور، سال ۱۳۸۸*.

سارانی، زینب، بهروز کشته‌گر و غلامرضا کشاورز حداد (۱۳۹۳)، «مشارکت زنان متأهل در بازار کار ایران: مدل‌سازی غیرخطی تابع لاجیت»، *فصلنامه مدل‌سازی اقتصادی*، سال هشتم، شماره ۳، صص ۱۳۴-۱۱۵.

کشاورز حداد، غلامرضا و مرتضی باقری قنبرآبادی (۱۳۹۰)، «تحلیل احتمال مشارکت زنان شهری و روستایی ایران در بازار کار با استفاده از روش‌های اقتصادسنجی پارامتریک و ناپارامتریک»، *مجله تحقیقات اقتصادی*، سال ششم، شماره ۹۶، صص ۱۷۴-۱۵۱.

کشاورز حداد، غلامرضا و آرش علویان قوانینی (۱۳۹۱)، «شکاف جنسیتی دستمزد در مناطق شهری ایران»، *فصلنامه پژوهش‌های اقتصادی ایران*، فصلنامه علمی-پژوهشکده علوم اقتصادی، سال هفدهم، شماره ۵۳، صص ۱۳۳-۱۰۱.

کولایی، الهه (۱۳۸۳)، «*جنسیت و توسعه در کشورهای خاورمیانه و شمال آفریقا*»، مرکز پژوهش‌های مجلس شورای اسلامی.

موحدی، بهروز (۱۳۸۱))، «چالش‌های اشتغال زنان در ایران»، *ماهنامه حقوق زنان*، سال چهارم، شماره ۲۱، صص ۸ -۳.

Bera, A. K., and Bilias, Y. (2001), "Rao's Score, Neyman's C (α) and Silvey's LM tests: An Essay on Historical Developments and Some New Results", *Journal of Statistical Planning and Inference*, vol. 97, pp. 9-44.

Berulava, G., and Chikava, G. (2012), "The Determinants of Household Labor Supply in Georgia, France and Romania: a Comparative Study". *Eurasian Journal of Business and Economics*, vol. 5, no.9, pp.141-164.

Bianco, A.M., and Martínez, E. (2009), "Robust Testing in the Logistic Regression Model", *Computational Statistics and Data Analysis*, vol. 53, pp. 4095-4105.

Carroll, R., Ruppert D., and Welsh, A. (1998), "Local Estimating Equations", *Journal of American Statistical Association*, vol. 93, pp. 214–27.

Chiappori, P.A. (1988), "Rational Household Labor Supply", *Econometrica*, vol. 56, pp. 63-89.

Chiappori, P.A., (1992), "Collective Labor Supply and Welfare", *Journal of Political Economy*, vol. 100, pp. 437-67.

Donni O., and Moreau, N. (2007), "Collective Labor Supply: a Single-Equation Model and Some Evidence from French Data", *The Journal of Human Resources*, vol. XLII, pp. 214-246.

Frolich, M. (2006), "Non-parametric Regression for Binary Dependent Variables", *Econometrics Journal*, vol. 9, pp. 511–540.

García, J., and María J. S. (2002), "*Female Labor Supply in Spain: The Importance of Behavioural Assumptions and Unobserved Heterogeneity Specification*", Department of Pampeu Fabra.

Gerfin, M. (1996), "Parametric and Semi-parametric Estimation of the Binary Response Model of Labor Market Participation", *Journal of Applied Econometrics*, vol. 11, pp. 321–39.

Gong, X., and van Soest, A. (2000), "Family Structure and Female Labour Supply in Mexico City", *IZA Discussion*, Paper No: 214.

Keshtegar, B., Piri. J., and Kisi, O. (2016), "A Nonlinear Mathematical Modeling of Daily Pan Evaporation based on Conjugate Gradient Method", *Computers and Electronics in Agriculture*, vol. 127, pp. 120-30.

Killingworth, M.R., and Heckman, J. (1986), "Female Labor Supply: a Survey*", In Handbook of Labor Economics*. Vol. 1.O. Ashenfelter and R. Layard eds. Amsterdam: North-Holland, pp. 101–204.

Klein, R. and Spady, R. (1993), "An Efficient Eemi-Parametric Estimator for Binary Response Models", *Econometrica*, vol. 6, no. 1, pp. 387–421.

Komarek, P. (2004), "*Logistic Regression for Data Mining and High-Dimensional Classification",* Ph.D. Thesis. Carnegie Mellon University.

Leckner, M. (1991), "Testing Logit Models in Practice", *Empirical Economics*, vol. 16, no. 2, pp. 177-198.

Maalouf, M., and Trafalis, T.B. (2011), "Robust Weighted kernel Logistic Regression in Imbalanced and Rare Events Data", *Computational Statistics and Data Analysis*, vol. 55, pp. 168-183.

Willmott, C.J., and Matsuura, K. (1998), "On the Use of Dimensioned Measures of Error to Evaluate the Performance of Spatial Interpolators", *International Journal of Geographical*, vol. 86, pp. 121-136.